The Molecular Biology Database Collection: 2005 update

Michael Y. Galperin*

National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

ABSTRACT

The *Nucleic Acids Research* Molecular Biology Database Collection is a public online resource that lists the databases described in this and previous issues of *Nucleic Acids Research* together with other databases of value to the biologist and available throughout the world. All databases included in this Collection are freely available to the public. The 2005 update includes 719 databases, 171 more than the 2004 one. The databases are organized in a hierarchical classification that simplifies the process of finding the right database for any given task. The growing number of databases related to immunology, plant and organelle research have been accommodated by separating them into three new categories. The database summaries provide brief descriptions of the databases, contact details, appropriate references and acknowledgements. The online summaries also serve as a venue for the maintainers of each database to introduce database updates and other improvements in the scope and tools. These updates are particularly important for those databases that have not been described in print in the recent past. The database list and summaries are available online at the *Nucleic Acids Research* web site, http://nar. oupjournals.org/.

COMMENTARY

In its 12th annual database issue, *Nucleic Acids Research* presents 135 new and recently updated molecular biology databases. The current release of the *Nucleic Acids Research* online Molecular Biology Database Collection (Table 1) includes 719 databases, an increase of 171 over last year (1). The database geography also continues to expand. This year we have the first databases from Brazil, Cuba, Estonia, Greece, Hungary (2,3), Malaysia, Taiwan (4,5) and Turkey.

The database authors have again shown remarkable creativity in naming their databases: last year's ORFanage [the database of orphan ORFs (6)] has been joined by H-ANGEL [Human ANatomic Gene Expression Library (7)], PROPHECY [PROfiling of PHEnotypic Characteristics in Yeast (8)], PANDIT [Protein and Associated Nucleotide Domains with Inferred Trees (9)], SIEGE [Smoking-Induced Epithelial Gene Expression (10)] and other aptly named databases. The database list is divided into 14 major categories, 3 more than last year. One of them, the category for immunology-related databases, was created in response to the rapid growth in databases dedicated to immuno-polymorphisms, certainly an offshoot of the Human Genome Project. The proliferation of plant-related databases, sparked by the completion of the first two plant genomes (*Arabidopsis thaliana* and *Oryza sativa*) and steady progress in sequencing other plants, prompted elevation of their status from a subcategory to a separate category. One more category, organelle databases, was created to provide a single home for the databases on chloroplasts and mitochondria from various sources. As always, we hope that these database listings, organized into a hierarchical structure, will help introduce the community of biologists to the enormous body of data accumulated by their colleagues and simplify the process of finding the appropriate database for each particular task.

Certainly, this listing is far from exhaustive. To be included, databases had to be publicly available to any user and allow direct browsing of the data without downloading any special software that might interfere with institutional firewalls. This means leaving out several potentially interesting database projects.

Of the 548 databases featured in last year's compilation, 17 have been dropped from the list because they have been discontinued, merged into larger ones or, like the well-known Kabat database, converted to commercial access. The previous year saw a loss of 13 databases from a total of 386 in the 2003 release. These numbers and the history of Swiss-Prot (http://www.expasy.org/announce/) and the GDB Human Genome Database (http://www.gdb.org/gdb/aboutGDB.html) show that the databases that offer useful content usually manage

Table 1. Molecular Biology Database Collection[a]

No.[b]	Database name	Full name and/or description	URL
	1. Nucleotide Sequence Databases		
	1.1. International Nucleotide Sequence Database Collaboration		
1	DDBJ—DNA Data Bank of Japan	All known nucleotide and protein sequences	http://www.ddbj.nig.ac.jp
2	EMBL Nucleotide Sequence Database	All known nucleotide and protein sequences	http://www.ebi.ac.uk/embl.html
3	GenBank®	All known nucleotide and protein sequences	http://www.ncbi.nlm.nih.gov/Entrez
	1.2. DNA sequences: genes, motifs and regulatory sites		
	1.2.1. Coding and coding DNA		
403	ACLAME	A classification of genetic mobile elements	http://aclame.ulb.ac.be/
30	CUTG	Codon usage tabulated from GenBank	http://www.kazusa.or.jp/codon/
480	Genetic Codes	Genetic codes in various organisms and organelles	http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi
668	Entrez Gene	Gene-centered information at NCBI	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
495	HERVd	Human endogenous retrovirus database	http://herv.img.cas.cz
687	Hoppsigen	Human and mouse homologous processed pseudogenes	http://pbil.univ-lyon1.fr/databases/hoppsigen.html
294	Imprinted Gene Catalogue	Imprinted genes and parent-of-origin effects in animals	http://www.otago.ac.nz/IGC
512	Islander	Pathogenicity islands and prophages in bacterial genomes	http://www.indiana.edu/~islander
343	MICdb	Prokaryotic microsatellites	http://www.cdfd.org.in/micas
707	NPRD	Nucleosome positioning region database	http://srs6.bionet.nsc.ru/srs6/
47	STRBase	Short tandem DNA repeats database	http://www.cstl.nist.gov/div831/strbase/
5	TIGR Gene Indices	Organism-specific databases of EST and gene sequences	http://www.tigr.org/tdb/tgi.shtml
48	Transterm	Codon usage, start and stop signals	http://uther.otago.ac.nz/Transterm.html
6	UniGene	Non-redundant set of eukaryotic gene-oriented clusters	http://www.ncbi.nlm.nih.gov/UniGene/
320	UniVec	Vector sequences, adapters, linkers and primers used in DNA cloning, can be used to check for vector contamination	http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html
302	VectorDB	Characterization and classification of nucleic acid vectors	http://genome-www2.stanford.edu/vectordb/
305	Xpro	Eukaryotic protein-encoding DNA sequences, both intron-containing and intron-less genes	http://origin.bic.nus.edu.sg/xpro/
	1.2.2. Gene structure, introns and exons, splice sites		
414	ASAP	Alternative spliced isoforms	http://www.bioinformatics.ucla.edu/ASAP
28	ASD	Alternative splicing database at EBI, includes three databases AltSplice, AltExtron and AEdb	http://www.ebi.ac.uk/asd
10	ASDB	Alternative splicing database: protein products and expression patterns of alternatively spliced genes	http://hazelton.lbl.gov/~teplitski/alt
639	ASHESdb	Alternatively spliced human genes by exon skipping database	http://sege.ntu.edu.sg/wester/ashes/
450	EASED	Extended alternatively spliced EST database	http://eased.bioinf.mdc-berlin.de/
667	ECgene	Genome annotation for alternative splicing	http://genome.ewha.ac.kr/ECgene/
631	EDAS	EST-derived alternative splicing database	http://www.ig-msk.ru:8005/EDAS/
34	ExInt	Exon–intron structure of eukaryotic genes	http://sege.ntu.edu.sg/wester/exint/
36	HS3D	*Homo sapiens* splice sites dataset	http://www.sci.unisannio.it/docenti/rampone/
238	Intronerator	Alternative splicing in *C.elegans* and *C.briggsae*	http://www.cse.ucsc.edu/~kent/intronerator/
46	SpliceDB	Canonical and non-canonical mammalian splice sites	http://www.softberry.com/berry.phtml?topic=splicedb&group=data&subgroup=spldb
746	SpliceInfo	Modes of alternative splicing in human genome	http://140.115.50.96/SpliceInfo/
580	SpliceNest	A tool for visualizing splicing of genes from EST data	http://splicenest.molgen.mpg.de/
	1.2.3. Transcriptional regulator sites and transcription factors		
231	ACTIVITY	Functional DNA/RNA site activity	http://wwwmgs.bionet.nsc.ru/mgs/systems/activity/
31	DBTBS	*Bacillus subtilis* promoters and transcription factors	http://dbtbs.hgc.jp/
663	DoOP	Database of orthologous promoters: chordates and plants	http://doop.abc.hu/
106	DPInteract	Binding sites for *E.coli* DNA-binding proteins	http://arep.med.harvard.edu/dpinteract
33	EPD	Eukaryotic promoter database	http://www.epd.isb-sib.ch

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
494	HemoPDB	Hematopoietic promoter database: transcriptional regulation in hematopoiesis	http://bioinformatics.med.ohio-state.edu/HemoPDB
516	JASPAR	PSSMs for transcription factor DNA-binding sites	http://jaspar.cgb.ki.se
700	MAPPER	Putative transcription factor binding sites in various genomes	http://bio.chip.org/mapper
40	PLACE	Plant cis-acting regulatory DNA elements	http://www.dna.affrc.go.jp/htdocs/PLACE
41	PlantCARE	Plant promoters and cis-acting regulatory elements	http://intra.psb.ugent.be:8080/PlantCARE/
563	PlantProm	Plant promoter sequences for RNA polymerase II	http://mendel.cs.rhul.ac.uk/
566	PRODORIC	Prokaryotic database of gene regulation networks	http://prodoric.tu-bs.de/
42	PromEC	E.coli promoters with experimentally identified transcriptional start sites	http://bioinfo.md.huji.ac.il/marg/promec
246	SELEX_DB	DNA and RNA binding sites for various proteins, found by systematic evolution of ligands by exponential enrichment	http://wwwmgs.bionet.nsc.ru/mgs/systems/selex/
227	TESS	Transcription element search system	http://www.cbil.upenn.edu/tess
756	TRACTOR db	Transcription factors in gamma-proteobacteria database	http://www.tractor.lncc.br/
345	TRANSCompel	Composite regulatory elements affecting gene transcription in eukaryotes	http://www.gene-regulation.com/pub/databases.html#transcompel
340	TRANSFAC	Transcription factors and binding sites	http://transfac.gbf.de/TRANSFAC/index.html
757	TRED	Transcriptional regulatory element database	http://rulai.cshl.edu/tred
49	TRRD	Transcription regulatory regions of eukaryotic genes	http://www.bionet.nsc.ru/trrd/
2. RNA sequence databases			
229	16S and 23S rRNA Mutation Database	16S and 23S ribosomal RNA mutations	http://www.fandm.edu/Departments/Biology/Databases/RNA.html
230	5S rRNA Database	5S rRNA sequences	http://biobases.ibch.poznan.pl/5SData/
411	Aptamer database	Small RNA/DNA molecules binding nucleic acids, proteins	http://aptamer.icmb.utexas.edu/
232	ARED	AU-rich element-containing mRNA database	http://rc.kfshrc.edu.sa/ared
378	Mobile group II introns	A database of group II introns, self-splicing catalytic RNAs	http://www.fp.ucalgary.ca/group2introns/
463	European rRNA database	All complete or nearly complete rRNA sequences	http://www.psb.ugent.be/rRNA/
490	GtRDB	Genomic tRNA database	http://rna.wustl.edu/GtRDB
236	Guide RNA Database	RNA editing in various kinetoplastid species	http://biosun.bio.tu-darmstadt.de/goringer/gRNA/gRNA.html
76	HIV Sequence Database	HIV RNA sequences	http://hiv-web.lanl.gov/
689	HuSiDa	Human siRNA database	http://itb1.biologie.hu-berlin.de/~nebulus/sirna/
237	HyPaLa	Hybrid pattern library: structural elements in classes of RNA	http://bibiserv.techfak.uni-bielefeld.de/HyPa/
379	IRESdb	Internal ribosome entry site database	http://ifr31w3.toulouse.inserm.fr/IRESdatabase/
529	microRNA Registry	Database of microRNAs (small non-coding RNAs)	http://www.sanger.ac.uk/Software/Rfam/mirna/
380	NCIR	Non-canonical interactions in RNA structures	http://prion.bchs.uh.edu/bp_type/
381	ncRNAs Database	Non-coding RNAs with regulatory functions	http://biobases.ibch.poznan.pl/ncRNA/
705	NONCODE	A database of non-coding RNAs	http://www.bioinfo.org.cn/NONCODE/index.htm
240	PLANTncRNAs	Plant non-coding RNAs	http://www.prl.msu.edu/PLANTncRNAs
564	Plant snoRNA DB	snoRNA genes in plant species	http://bioinf.scri.sari.ac.uk/cgi-bin/plant_snorna/home
723	PolyA_DB	A database of mammalian mRNA polyadenylation	http://polya.umdnj.edu/polyadb/
242	PseudoBase	Database of RNA pseudoknots	http://wwwbio.leidenuniv.nl/~Batenburg/PKB.html
382	Rfam	Non-coding RNA families	http://www.sanger.ac.uk/Software/Rfam/
244	RISSC	Ribosomal internal spacer sequence collection	http://ulises.umh.es/RISSC
630	RNAdb	Mammalian non-coding RNA database	http://ncrna.bioinformatics.com.au/
245	RNA Modification Database	Naturally modified nucleosides in RNA	http://medlib.med.utah.edu/RNAmods/
43	RRNDB	rRNA operon numbers in various prokaryotes	http://rrndb.cme.msu.edu/
629	siRNAdb	siRNA database and search engine	http://sirna.cgb.ki.se/
247	Small RNA Database	Small RNAs from prokaryotes and eukaryotes	http://mbcr.bcm.tmc.edu/smallRNA
248	SRPDB	Signal recognition particle database	http://psyche.uthct.edu/dbs/SRPDB/SRPDB.html
754	SSU rRNA Modification Database	Modified nucleosides in small subunit rRNA	http://medstat.med.utah.edu/SSUmods/
383	Subviral RNA Database	Viroids and viroid-like RNAs	http://subviral.med.uottawa.ca/
249	tmRNA Website	tmRNA sequences and alignments	http://www.indiana.edu/~tmrna
250	tmRDB	tmRNA database	http://psyche.uthct.edu/dbs/tmRDB/tmRDB.html
251	tRNA sequences	tRNA viewer and sequence editor	http://www.uni-bayreuth.de/departments/biochemie/trna/
252	UTRdb/UTRsite	5'- and 3'-UTRs of eukaryotic mRNAs	http://bighost.area.ba.cnr.it/srs6/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
	3. Protein sequence databases		
	3.1. General sequence databases		
163	EXProt	Sequences of proteins with experimentally verified function	http://www.cmbi.kun.nl/EXProt/
542	NCBI Protein database	All protein sequences: translated from GenBank and imported from other protein databases	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein
714	PA-GOSUB	Protein sequences from model organisms, GO assignment and subcellular localization	http://www.cs.ualberta.ca/~bioinfo/PA/ProteomeIndex.html
194	PIR-PSD	Protein information resource protein sequence database, has been merged into the UniProt knowledgebase	http://pir.georgetown.edu/
370	PIR-NREF	PIR's non-redundant reference protein database	http://pir.georgetown.edu/pirwww/pirnref.shtml
565	PRF	Protein research foundation database of peptides: sequences, literature and unnatural amino acids	http://www.prf.or.jp/en
197	Swiss-Prot	Now UniProt/Swiss-Prot: expertly curated protein sequence database, section of the UniProt knowledgebase	http://www.expasy.org/sprot
198	TrEMBL	Now UniProt/TrEMBL: computer-annotated translations of EMBL nucleotide sequence entries: section of the UniProt knowledgebase	http://www.expasy.org/sprot
775	UniParc	UniProt archive: a repository of all protein sequences, consisting only of unique identifiers and sequence	http://www.uniprot.org/database/archive.shtml
318	UniProt	Universal protein knowledgebase: merged data from Swiss-Prot, TrEMBL and PIR protein sequence databases	http://www.uniprot.org/
776	UniRef	UniProt non-redundant reference database: clustered sets of related sequences (including splice variants and isoforms)	http://www.uniprot.org/database/nref.shtml
	3.2. Protein properties		
221	AAindex	Physicochemical properties of amino acids	http://www.genome.ad.jp/aaindex/
729	ProNIT	Thermodynamic data on protein–nucleic acid interactions	http://gibk26.bse.kyutech.ac.jp/jouhou/pronit/pronit.html
280	ProTherm	Thermodynamic data for wild-type and mutant proteins	http://gibk26.bse.kyutech.ac.jp/jouhou/protherm/protherm.html
772	TECRdb	Thermodynamics of enzyme-catalyzed reactions	http://xpdb.nist.gov/enzyme_thermodynamics/
	3.3. Protein localization and targeting		
444	DBSubLoc	Database of protein subcellular localization	http://www.bioinfo.tsinghua.edu.cn/dbsubloc.html
375	NESbase	Nuclear export signals database	http://www.cbs.dtu.dk/databases/NESbase
376	NLSdb	Nuclear localization signals	http://cubic.bioc.columbia.edu/db/NLSdb/
704	NMPdb	Nuclear matrix associated proteins database	http://www.rostlab.org/db/NMPdb/
706	NOPdb	Nucleolar proteome database	http://www.lamondlab.com/NOPdb/
734	PSORTdb	Protein subcellular localization in bacteria	http://db.psort.org/
745	SPD	Secreted protein database	http://spd.cbi.pku.edu.cn
587	THGS	Transmembrane helices in genome sequences	http://pranag.physics.iisc.ernet.in/thgs/
589	TMPDB	Experimentally characterized transmembrane topologies	http://bioinfo.si.hirosaki-u.ac.jp/~TMPDB/
	3.4. Protein sequence motifs and active sites		
374	ASC	Active sequence collection: biologically active peptides	http://bioinformatica.isa.cnr.it/ASC/
203	Blocks	Alignments of conserved regions in protein families	http://blocks.fhcrc.org/
440	CSA	Catalytic site atlas: active sites and catalytic residues in enzymes of known 3D structure	http://www.ebi.ac.uk/thornton-srv/databases/CSA/
438	COMe	Co-ordination of metals etc.: classification of bioinorganic proteins (metalloproteins and some other complex proteins)	http://www.ebi.ac.uk/come
771	CopS	Comprehensive peptide signature database	http://203.195.151.46/copsv2/index.html
666	eBLOCKS	Highly conserved protein sequence blocks	http://fold.stanford.edu/eblocks/acsearch.html
206	eMOTIF	Protein sequence motif determination and searches	http://motif.stanford.edu/emotif
179	Metalloprotein Site Database	Metal-binding sites in metalloproteins	http://metallo.scripps.edu/
209	O-GlycBase	O- and C-linked glycosylation sites in proteins	http://www.cbs.dtu.dk/databases/OGLYCBASE/
717	PDBSite	3D structure of protein functional sites	http://srs6.bionet.nsc.ru/srs6/
187	Phospho.ELM	S/T/Y protein phosphorylation sites (formerly PhosphoBase)	http://phospho.elm.eu.org/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
193	PROMISE	Prosthetic centers and metal ions in protein active sites	http://metallo.scripps.edu/PROMISE
215	PROSITE	Biologically significant protein patterns and profiles	http://www.expasy.org/prosite
732	ProTeus	Signature sequences at the protein N- and C-termini	http://www.proteus.cs.huji.ac.il/
	3.5. Protein domain databases; protein classification		
622	ADDA	A database of protein domain classification	http://ekhidna.biocenter.helsinki.fi:8080/examples/servlets/adda/
204	CDD	Conserved domain database; includes protein domains from Pfam, SMART, COG and KOG databases	http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml
205	CluSTr	Clusters of Swiss-Prot+TrEMBL proteins	http://www.ebi.ac.uk/clustr
671	FunShift	Functional divergence between the subfamilies of a protein domain family	http://funshift.cgb.ki.se/
200	Hits	A database of protein domains and motifs	http://hits.isb-sib.ch/
207	InterPro	Integrated resource of protein families, domains and functional sites	http://www.ebi.ac.uk/interpro
208	iProClass	Integrated protein classification database	http://pir.georgetown.edu/iproclass/
561	PIRSF	Family/superfamily classification of whole proteins	http://pir.georgetown.edu/pirsf/
212	PRINTS	Hierarchical gene family fingerprints	http://www.bioinf.man.ac.uk/dbbrowser/PRINTS/
210	Pfam	Protein families: multiple sequence alignments and profile hidden Markov models of protein domains	http://www.sanger.ac.uk/Software/Pfam/
727	PRECISE	Predicted and consensus interaction sites in enzymes	http://precise.bu.edu/precisedb/
214	ProDom	Protein domain families	http://www.toulouse.inra.fr/prodom.html
216	ProtoMap	Hierarchical classification of Swiss-Prot proteins	http://protomap.cornell.edu/
567	ProtoNet	Hierarchical clustering of Swiss-Prot proteins	http://www.protonet.cs.huji.ac.il/
740	S4	Structure-based sequence alignments of SCOP superfamilies	http://compbio.mds.qmw.ac.uk/~james/S4.shtml
217	SBASE	Protein domain sequences and tools	http://www.icgeb.org/sbase
218	SMART	Simple modular architecture research tool: signalling, extracellular and chromatin-associated protein domains	http://smart.embl-heidelberg.de/
219	SUPFAM	Grouping of sequence families into superfamilies	http://pauling.mbu.iisc.ernet.in/~supfam
220	SYSTERS	Systematic re-searching and clustering of proteins	http://systers.molgen.mpg.de/
199	TIGRFAMs	TIGR protein families adapted for functional annotation	http://www.tigr.org/TIGRFAMs
	3.6. Databases of individual protein families		
156	AARSDB	Aminoacyl-tRNA synthetase database	http://rose.man.poznan.pl/aars/index.html
308	ASPD	Artificial selected proteins/peptides database	http://wwwmgs.bionet.nsc.ru/mgs/gnw/aspd/
158	BacTregulators	Transcriptional regulators of AraC and TetR families	http://www.bactregulators.org/
364	CSDBase	Cold shock domain-containing proteins	http://www.chemie.uni-marburg.de/~csdbase/
653	CuticleDB	Structural proteins of Arthropod cuticle	http://bioinformatics.biol.uoa.gr/cuticleDB
658	DCCP	Database of copper-chelating proteins	http://sdbi.sdut.edu.cn/DCCP/en/index.php
160	DExH/D Family Database	DEAD-box, DEAH-box and DExH-box proteins	http://www.helicase.net/dexhd/dbhome.htm
161	Endogenous GPCR List	G protein-coupled receptors; expression in cell lines	http://www.tumor-gene.org/GPCR/gpcr.html
162	ESTHER	Esterases and other alpha/beta hydrolase enzymes	http://www.ensam.inra.fr/esther
464	EyeSite	Families of proteins functioning in the eye	http://eyesite.cryst.bbk.ac.uk/
166	GPCRDB	G protein-coupled receptors database	http://www.gpcr.org/7tm/
679	gpDB	G-proteins and their interaction with GPCRs	http://bioinformatics.biol.uoa.gr/gpDB
167	Histone Database	Histone fold sequences and structures	http://research.nhgri.nih.gov/histones/
169	Homeobox Page	Homeobox proteins, classification and evolution	http://www.biosci.ki.se/groups/tbu/homeo.html
293	Hox-Pro	Homeobox genes database	http://www.iephb.nw.ru/labs/lab38/spirov/hox_pro/hox-pro00.html
170	Homeodomain Resource	Homeodomain sequences, structures and related genetic and genomic information	http://research.nhgri.nih.gov/homeodomain/
366	HORDE	Human olfactory receptor data exploratorium	http://bioinfo.weizmann.ac.il/HORDE/
174	InBase	Inteins (protein splicing elements) database: properties, sequences, bibliography	http://www.neb.com/neb/inteins.html
518	KinG—Kinases in Genomes	S/T/Y-specific protein kinases encoded in complete genomes	http://hodgkin.mbu.iisc.ernet.in/~king

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
519	Knottins	Database of knottins—small proteins with an unusual 'disulfide through disulfide' knot	http://knottin.cbs.cnrs.fr
176	LGICdb	Ligand-gated ion channel subunit sequences database	http://www.pasteur.fr/recherche/banques/LGIC/LGIC.html
368	Lipase Engineering Database	Sequence, structure and function of lipases and esterases	http://www.led.uni-stuttgart.de/
524	LOX-DB	Mammalian, invertebrate, plant and fungal lipoxygenases	http://www.dkfz-heidelberg.de/spec/lox-db/
177	MEROPS	Database of proteolytic enzymes (peptidases)	http://merops.sanger.ac.uk/
369	NPD	Nuclear protein database	http://npd.hgu.mrc.ac.uk/
546	NucleaRDB	Nuclear receptor superfamily	http://www.receptors.org/NR/
182	Nuclear Receptor Resource	Nuclear receptor superfamily	http://nrr.georgetown.edu/NRR/nrrhome.htm
183	NUREBASE	Nuclear hormone receptors database	http://www.ens-lyon.fr/LBMC/laudet/nurebase.html
184	Olfactory Receptor Database	Sequences for olfactory receptor-like molecules	http://senselab.med.yale.edu/senselab/ordb/
185	ooTFD	Object-oriented transcription factors database	http://www.ifti.org/ootfd
188	PKR	Protein kinase resource: sequences, enzymology, genetics and molecular and structural properties	http://pkr.sdsc.edu/html/index.shtml
759	PLPMDB	Pyridoxal-5'-phosphate dependent enzymes mutations	http://www.studiofmp.com/plpmdb/
609	ProLysED	A database of bacterial protease systems	http://genome.ukm.my/prolyses/
192	Prolysis	Proteases and natural and synthetic protease inhibitors	http://delphi.phys.univ-tours.fr/Prolysis/
224	REBASE	Restriction enzymes and associated methylases	http://rebase.neb.com/rebase/rebase.html
195	Ribonuclease P Database	RNase P sequences, alignments and structures	http://www.mbio.ncsu.edu/RNaseP/home.html
573	RPG	Ribosomal protein gene database	http://ribosome.miyazaki-med.ac.jp/
575	RTKdb	Receptor tyrosine kinase sequences	http://pbil.univ-lyon1.fr/RTKdb/
309	S/MARt dB	Nuclear scaffold/matrix attached regions	http://smartdb.bioinf.med.uni-goettingen.de/
741	Scorpion	Database of scorpion toxins	http://research.i2r.a-star.edu.sg:8080/scorpion/
372	SDAP	Structural database of allergenic proteins and food allergens	http://fermi.utmb.edu/SDAP
196	SENTRA	Sensory signal transduction proteins	http://www-wit.mcs.anl.gov/sentra/
373	SEVENS	7-transmembrane helix receptors (G-protein-coupled)	http://sevens.cbrc.jp/
248	SRPDB	Proteins of the signal recognition particles	http://bio.lundberg.gu.se/dbs/SRPDB/SRPDB.html
314	TrSDB	Transcription factor database	http://ibb.uab.es/trsdb
399	VKCDB	Voltage-gated potassium channel database	http://vkcdb.biology.ualberta.ca/
202	Wnt Database	Wnt proteins and phenotypes	http://www.stanford.edu/~rnusse/wntwindow.html
	4. Structure Databases		
	4.1. Small molecules		
646	ChEBI	Chemical entities of biological interest	http://www.ebi.ac.uk/chebi/
261	CSD	Cambridge structural database: crystal structure information for organic and metal-organic compounds	http://www.ccdc.cam.ac.uk/prods/csd/csd.html
265	HIC-Up	Hetero-compound Information Centre—Uppsala	http://xray.bmc.uu.se/hicup
402	AANT	Amino acid–nucleotide interaction database	http://aant.icmb.utexas.edu/
111	Klotho	Collection and categorization of biological compounds	http://www.biocheminfo.org/klotho
113	LIGAND	Chemical compounds and reactions in biological pathways	http://www.genome.ad.jp/ligand/
615	PDB-Ligand	3D structures of small molecules bound to proteins and nucleic acids	http://www.idrtech.com/PDB-Ligand/
735	PubChem	Structures and biological activities of small organic molecules	http://pubchem.ncbi.nlm.nih.gov/
	4.2. Carbohydrates		
429	CCSD	Complex carbohydrate structure database (CarbBank)	http://bssv01.lancs.ac.uk/gig/pages/gag/carbbank.htm
652	CSS	Carbohydrate structure suite: carbohydrate 3D structures derived from the PDB	http://www.dkfz.de/spec/css/
486	Glycan	Carbohydrate database, part of the KEGG system	http://glycan.genome.ad.jp/
292	GlycoSuiteDB	N- and O-linked glycan structures and biological sources	http://www.glycosuite.com/
535	Monosaccharide Browser	Space-filling Fischer projections of monosaccharides	http://www.jonmaber.demon.co.uk/monosaccharide
300	SWEET-DB	Annotated carbohydrate structure and substance information	http://www.dkfz-heidelberg.de/spec2/sweetdb/
	4.3. Nucleic acid structure		
272	NDB	Nucleic acid-containing structures	http://ndbserver.rutgers.edu/
273	NTDB	Thermodynamic data for nucleic acids	http://ntdb.chem.cuhk.edu.hk/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
387	RNABase	RNA-containing structures from PDB and NDB	http://www.rnabase.org/
283	SCOR	Structural classification of RNA: RNA motifs by structure, function and tertiary interactions	http://scor.lbl.gov/
	4.4. Protein structure		
413	ArchDB	Automated classification of protein loop structures	http://gurion.imim.es/archdb
255	ASTRAL	Sequences of domains of known structure, selected subsets and sequence–structure correspondences	http://astral.stanford.edu/
288	BAliBASE	A database for comparison of multiple sequence alignments	http://www-igbmc.u-strasbg.fr/BioInfo/BAliBASE2/index.html
257	BioMagResBank	NMR spectroscopic data for proteins and nucleic acids	http://www.bmrb.wisc.edu/
384	CADB	Conformational angles in proteins database	http://cluster.physics.iisc.ernet.in/cadb/
258	CATH	Protein domain structures database	http://www.biochem.ucl.ac.uk/bsm/cath_new
259	CE	3D protein structure alignments	http://cl.sdsc.edu/ce.html
260	CKAAPs DB	Structurally similar proteins with dissimilar sequences	http://ckaap.sdsc.edu/
442	Dali	Protein fold classification using the Dali search engine	http://www.bioinfo.biocenter.helsinki.fi:8080/dali/
385	Decoys 'R' Us	Computer-generated protein conformations	http://dd.stanford.edu/
447	DisProt	Database of Protein Disorder: proteins that lack fixed 3D structure in their native states	http://divac.ist.temple.edu/disprot
448	DomIns	Domain insertions in known protein structures	http://stash.mrc-lmb.cam.ac.uk/DomIns
264	DSDBASE	Native and modeled disulfide bonds in proteins	http://www.ncbs.res.in/~faculty/mini/dsdbase/dsdbase.html
386	DSMM	Database of simulated molecular motions	http://projects.villa-bosch.de/dbase/dsmm/
452	eF-site	Electrostatic surface of Functional site: electrostatic potentials and hydrophobic properties of the active sites	http://ef-site.protein.osaka-u.ac.jp/eF-site
674	GenDiS	Genomic distribution of protein structural superfamilies	http://caps.ncbs.res.in/gendis/home.html
472	Gene3D	Precalculated structural assignments for whole genomes	http://www.biochem.ucl.ac.uk/bsm/cath_new/Gene3D/
489	GTD	Genomic threading database: structural annotations of complete proteomes	http://bioinf.cs.ucl.ac.uk/GTD
322	GTOP	Protein fold predictions from genome sequences	http://spock.genes.nig.ac.jp/~genome/
360	Het-PDB Navi	Hetero-atoms in protein structures	http://daisy.nagahama-i-bio.ac.jp/golab/hetpdbnavi.html
498	HOMSTRAD	Homologous structure alignment database: curated structure-based alignments for protein families	http://www-cryst.bioc.cam.ac.uk/homstrad
267	IMB Jena Image Library	Visualization and analysis of 3D biopolymer structures	http://www.imb-jena.de/IMAGE.html
502	IMGT/3Dstructure-DB	Sequences and 3D structures of vertebrate immunoglobulins, T cell receptors and MHC proteins	http://imgt3d.igh.cnrs.fr/
268	ISSD	Integrated sequence–structure database	http://www.protein.bio.msu.su/issd
269	LPFC	Library of protein family core structures	http://www-smi.stanford.edu/projects/helix/LPFC
270	MMDB	NCBI's database of 3D structures, part of NCBI Entrez	http://www.ncbi.nlm.nih.gov/Structure
456	E-MSD	EBI's macromolecular structure database	http://www.ebi.ac.uk/msd
331	ModBase	Annotated comparative protein structure models	http://alto.compbio.ucsf.edu/modbase-cgi/index.cgi
262	MolMovDB	Database of macromolecular movements: descriptions of protein and macromolecular motions, including movies	http://bioinfo.mbb.yale.edu/MolMovDB/
274	PALI	Phylogeny and alignment of homologous protein structures	http://pauling.mbu.iisc.ernet.in/~pali
275	PASS2	Structural motifs of protein superfamilies	http://ncbs.res.in/~faculty/mini/campass/pass.html
557	PepConfDB	A database of peptide conformations	http://www.peptidome.org/products/list.htm
276	PDB	Protein structure databank: all publicly available 3D structures of proteins and nucleic acids	http://www.rcsb.org/pdb
277	PDB-REPRDB	Representative protein chains, based on PDB entries	http://mbs.cbrc.jp/pdbreprdb-cgi/reprdb_menu.pl
278	PDBsum	Summaries and analyses of PDB structures	http://www.biochem.ucl.ac.uk/bsm/pdbsum
619	PDB_TM	Transmembrane proteins with known 3D structure	http://www.enzim.hu/PDB_TM/
719	Protein Folding Database	Experimental data on protein folding	http://pfd.med.monash.edu.au
282	SCOP	Structural classification of proteins	http://scop.mrc-lmb.cam.ac.uk/scop
284	Sloop	Classification of protein loops	http://www-cryst.bioc.cam.ac.uk/~sloop/
583	Structure Superposition Database	Pairwise superposition of TIM-barrel structures	http://ssd.rbvi.ucsf.edu/
585	SWISS-MODEL Repository	Database of annotated 3D protein structure models	http://swissmodel.expasy.org/repository

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
285	SUPERFAMILY	Assignments of proteins to structural superfamilies	http://supfam.org/
584	SURFACE	Surface residues and functions annotated, compared and evaluated: a database of protein surface patches	http://cbm.bio.uniroma2.it/surface
764	TargetDB	Target data from worldwide structural genomics projects	http://targetdb.pdb.org/
401	3D-GENOMICS	Structural annotations for complete proteomes	http://www.sbg.bio.ic.ac.uk/3dgenomics
310	TOPS	Topology of protein structures database	http://www.tops.leeds.ac.uk
	5. Genomics Databases (non-human)		
	5.1. Genome annotation terms, ontologies and nomenclature		
73	Genew	Human gene nomenclature: approved gene symbols	http://www.gene.ucl.ac.uk/nomenclature
487	GO	Gene ontology consortium database	http://www.geneontology.org/
389	GOA	EBI's gene ontology annotation project	http://www.ebi.ac.uk/GOA
513	IUBMB Nomenclature database	Nomenclature of enzymes, membrane transporters, electron transport proteins and other proteins	http://www.chem.qmul.ac.uk/iubmb
514	IUPAC Nomenclature database	Nomenclature of biochemical and organic compounds approved by the IUBMB-IUPAC Joint Commission	http://www.chem.qmul.ac.uk/iupac
515	IUPHAR-RD	The International Union of Pharmacology recommendations on receptor nomenclature and drug classification	http://www.iuphar-db.org/iuphar-rd/
552	PANTHER	Gene products organized by biological function	http://panther.celera.com/
317	UMLS	Unified medical language system	http://umlsks.nlm.nih.gov/
	5.1.1. Taxonomy and Identification		
78	ICB	*gyrB* database for identification and classification of bacteria	http://seasquirt.mbio.co.jp/icb/index.php
297	NCBI Taxonomy	Names of all organisms represented in GenBank	http://www.ncbi.nlm.nih.gov/Taxonomy/
608	PANDIT	Protein and associated nucleotide domains with inferred trees	http://www.ebi.ac.uk/goldman-srv/pandit/
299	RIDOM	rRNA-based differentiation of medical microorganisms	http://www.ridom-rdna.de/
243	RDP-II	Ribosomal database project	http://rdp.cme.msu.edu
301	Tree of Life	Information on phylogeny and biodiversity	http://phylogeny.arizona.edu/tree/phylogeny.html
	5.2. General genomics databases		
7	COG	Clusters of orthologous groups of proteins	http://www.ncbi.nlm.nih.gov/COG
650	COGENT	Complete genome tracking; predicted peptides from fully sequenced genomes	http://maine.ebi.ac.uk:8000/services/cogent/
337	CORG	Comparative regulatory genomics: conserved non-coding sequence blocks	http://corg.molgen.mpg.de/
445	DEG	Database of essential genes from bacteria and yeast	http://tubic.tju.edu.cn/deg
451	EBI Genomes	EBI's collection of databases for the analysis of complete and unfinished viral, pro- and eukaryotic genomes	http://www.ebi.ac.uk/genomes
453	EGO	Eukaryotic gene orthologs: orthologous DNA sequences in the TIGR gene indices	http://www.tigr.org/tdb/tgi/ego/
70	EMGlib	Enhanced microbial genomes library: completely sequenced genomes of unicellular organisms	http://pbil.univ-lyon1.fr/emglib/emglib.html
458	Entrez Genomes	NCBI's collection of databases for the analysis of complete and unfinished viral, pro- and eukaryotic genomes	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Genome
461	ERGOLight	Integrated biochemical data on nine bacterial genomes: publicly available portion of the ERGO database	http://www.ergo-light.com/ERGO
470	FusionDB	Database of bacterial and archaeal gene fusion events	http://igs-server.cnrs-mrs.fr/FusionDB
611	Genome Atlas	DNA structural properties of sequenced genomes	http://www.cbs.dtu.dk/services/GenomeAtlas/
484	Genome Information Broker	DDBJ's collection of databases for the analysis of complete and unfinished viral, pro- and eukaryotic genomes	http://gib.genes.nig.ac.jp
678	Genome Reviews	Integrated view of complete genomes	http://www.ebi.ac.uk/GenomeReviews/
75	GOLD	Genomes online database: a listing of completed and ongoing genome projects	http://www.genomesonline.org/
352	HGT-DB	Putative horizontally transferred genes in prokaryotic genomes	http://www.fut.es/~debb/HGT/
223	Integr8	Functional classification of proteins in whole genomes	http://www.ebi.ac.uk/integr8/

Table 1. Continued

No.[b]	Database name	Full name and/or description	URL
112	KEGG	Kyoto encyclopedia of genes and genomes: integrated suite of databases on genes, proteins and metabolic pathways	http://www.genome.jp/kegg
528	MBGD	Microbial genome database for comparative analysis	http://mbgd.genome.ad.jp/
549	ORFanage	Database of orphan ORFs (ORFs with no homologs) in complete microbial genomes	http://www.cs.bgu.ac.il/~nomsiew/ORFans
551	PACRAT	Archaeal and bacterial intergenic sequence features	http://www.biosci.ohio-tate.edu/~pacrat
715	PartiGeneDB	Assembled partial genomes for ~250 eukaryotic organisms	http://www.partigenedb.org/
354	PEDANT	Results of an automated analysis of genomic sequences	http://pedant.gsf.de/
99	TIGR Microbial Database	Lists of completed and ongoing genome projects with links to complete genome sequences	http://www.tigr.org/tdb/mdb/mdbcomplete.html
66	TIGR Comprehensive Microbial Resource	Various data on complete microbial genomes: uniform annotation, properties of DNA and predicted proteins	http://www.tigr.org/CMR
311	TransportDB	Predicted membrane transporters in complete genomes, classified according to the TC classification system	http://www.membranetransport.org/
118	WIT3	What is there? Metabolic reconstruction for completely sequenced microbial genomes	http://www-wit.mcs.anl.gov/wit3/
	5.3. Organism-specific databases		
	5.3.1. Viruses		
473	HCVDB	The hepatitis C virus database	http://hepatitis.ibcp.fr/
497	HIV Drug Resistance Database	HIV mutations that confer resistance to anti-HIV drugs	http://resdb.lanl.gov/Resist_DB/default.htm
168	HIV Molecular Immunology Database	HIV epitopes	http://hiv-web.lanl.gov/immunology/
365	HIV RT and Protease Sequence Database	HIV reverse transcriptase and protease sequences	http://hivdb.stanford.edu/
602	NCBI Viral Genomes	Viral genome resource at NCBI	http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/viruses.html
725	Poxvirus.org	Poxvirus genomic sequences and gene annotation	http://www.poxvirus.org/
750	T4-like genome database	Sequences of T4-like bacteriophages from various sources	http://phage.bioc.tulane.edu/
201	VIDA	Homologous viral protein families database	http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html
761	VIPER	Virus particle explorer: virus capsid structures	http://mmtsb.scripps.edu/viper
303	VirOligo	Virus-specific oligonucleotides for PCR and hybridization	http://viroligo.okstate.edu/
	5.3.2. Prokaryotes		
641	BacMap	Picture atlas of annotated bacterial genomes	http://wishart.biology.ualberta.ca/BacMap
614	MetaGrowth	Growth requirements of bacterial pathogens	http://igs-server.cnrs-mrs.fr/axenic/
720	PGTdb	Prokaryotic growth temperature database	http://pgtdb.csie.ncu.edu.tw/
	5.3.2.1. Escherichia coli		
415	ASAP	A systematic annotation package for community analysis of E.coli and related genomes	https://asap.ahabs.wisc.edu/annotation/php/ASAP1.htm
428	CyberCell database	A collection of data on E.coli K12 intended for mathematical modeling to simulate the bacterial cell	http://redpoll.pharmacy.ualberta.ca/CCDB
436	coliBase	A database for E.coli, Salmonella and Shigella	http://colibase.bham.ac.uk/
437	Colibri	E.coli genome database at Institut Pasteur	http://genolist.pasteur.fr/Colibri/
623	EchoBASE	Post-genomic studies of Escherichia coli	http://www.ecoli-york.org/
462	Essential genes in E.coli	First results of an E.coli gene deletion project	http://www.genome.wisc.edu/resources/essential.htm
482	GenoBase	E.coli genome database at Nara Institute	http://ecoli.aist-nara.ac.jp/
165	GenProtEC	E.coli K12 genome and proteome database	http://genprotec.mbl.edu
555	PEC	Profiling of E.coli chromosome	http://shigen.lab.nig.ac.jp/ecoli/pec
108	EcoCyc	E.coli K12 genes, metabolic pathways, transporters and gene regulation	http://ecocyc.org/
69	EcoGene	Sequence and literature data on E.coli genes and proteins	http://bmb.med.miami.edu/EcoGene/EcoWeb/
116	RegulonDB	Transcriptional regulation and operon organization in E.coli	http://www.cifn.unam.mx/Computational_Genomics/regulondb/
	5.3.2.2. Bacillus subtilis		
424	BSORF	Bacillus subtilis genome database at Kyoto U.	http://bacillus.genome.ad.jp/
89	NRSub	Non-redundant Bacillus subtilis database at U. Lyon	http://pbil.univ-lyon1.fr/nrsub/nrsub.html
96	SubtiList	Bacillus subtilis genome database at Institut Pasteur	http://genolist.pasteur.fr/SubtiList/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
	5.3.2.3. Other bacteria		
420	BioCyc	Pathway/genome databases for many bacteria	http://biocyc.org/
426	CampyDB	Database for *Campylobacter* genome analysis	http://campy.bham.ac.uk/
433	ClostriDB	Finished and unfinished genomes of *Clostridium* spp.	http://clostri.bham.ac.uk/
648	CIDB	*Chlamydia* Interactive Database: gene expression data	http://www.it.deakin.edu.au/CIDB
68	CyanoBase	Cyanobacterial genomes	http://www.kazusa.or.jp/cyano
521	LeptoList	*Leptospira interrogans* genome	http://bioinfo.hku.hk/LeptoList
534	MolliGen	Genomic data on mollicutes	http://cbi.labri.fr/outils/molligen/
733	PseudoCAP	*Pseudomonas aeruginosa* genome database and community annotation project	http://www.pseudomonas.com/
94	RsGDB	*Rhodobacter sphaeroides* genome	http://www.hgsc.bcm.tmc.edu/projects/microbial/Rsphaeroides/
762	VirFact	Bacterial virulence factors and pathogenicity islands	http://virfact.burnham.org/
760	Virulence Factors	Reference database for microbial virulence factors	http://zdsys.chgb.org.cn/VFs/main.htm
	5.3.3. Unicellular eukaryotes		
409	ApiEST-DB	EST sequences from various *Apicomplexan* parasites	http://www.cbil.upenn.edu/paradbs-servlet
439	CryptoDB	*Cryptosporidium parvum* genome database	http://cryptodb.org/
662	Diatom EST Database	ESTs from two diatom algae, *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*	http://avesthagen.sznbowler.com/
446	DictyBase	Universal resource for *Dictyostelium discoideum*	http://dictybase.org/
72	Full-Malaria	Full-length cDNA library from erythrocytic-stage *Plasmodium falciparum*	http://fullmal.ims.u-tokyo.ac.jp/
328	GeneDB	Curated database for various Sanger-sequenced genomes	http://www.genedb.org/
698	LumbriBASE	ESTs of the earthworm *Lumbricus rubellus*	http://www.earthworms.org/
91	PlasmoDB	*Plasmodium* genome database	http://plasmodb.org/
586	TcruziDB	*Trypanosoma cruzi* genome database	http://tcruzidb.org/
359	ToxoDB	*Toxoplasma gondii* genome database	http://toxodb.org/
	5.3.4. Fungi		
	5.3.4.1. Yeasts		
635	AGD	*Ashbya gossypii* genome database	http://agd.unibas.ch
617	CandidaDB	*Candida albicans* genome database	http://genolist.pasteur.fr/CandidaDB
645	Candida Genome	*Candida albicans* genome database	http://www.candidagenome.org/
441	CYGD	MIPS Comprehensive yeast genome database	http://mips.gsf.de/proj/yeast
483	Génolevures	A comparison of *S.cerevisiae* and 14 other yeast species	http://cbi.labri.fr/Genolevures
730	PROPHECY	Profiling of phenotypic characteristics in yeast	http://prophecy.lundberg.gu.se/
576	SCMD	*Saccharomyces cerevisiae* morphological database: micrographs of budding yeast mutants	http://yeast.gi.k.u-tokyo.ac.jp/
577	SCPD	*Saccharomyces cerevisiae* promoter database	http://cgsigma.cshl.org/jian
357	SGD	*Saccharomyces* genome database	http://www.yeastgenome.org/
25	TRIPLES	Transposon-insertion phenotypes, localization and expression in *Saccharomyces*	http://ygac.med.yale.edu/triples/
306	YDPM	Yeast deletion project and mitochondria database	http://www-deletion.stanford.edu/YDPM/YDPM_index.html
342	Yeast Intron Database	Ares lab database of spliceosomal introns in *S.cerevisiae*	http://www.cse.ucsc.edu/research/compbio/yeast_introns.html
254	Yeast snoRNA Database	Yeast small nucleolar RNAs	http://www.bio.umass.edu/biochem/rna-sequence/Yeast_snoRNA_Database/snoRNA_DataBase.html
307	yMGV	Yeast microarray global viewer	http://www.transcriptome.ens.fr/ymgv/
763	YRC PDR	Yeast resource center public data repository	http://www.yeastrc.org/pdr/
	5.3.4.2. Other fungi		
425	CADRE	Central *Aspergillus* data repository	http://www.cadre.man.ac.uk/
435	COGEME	Phytopathogenic fungi and oomycete EST database	http://cogeme.ex.ac.uk
533	MNCDB	MIPS *Neurospora crassa* database	http://mips.gsf.de/proj/neurospora/
708	OGD	Oomycete Genomics Database: ESTs and annotation	http://www.oomycete.net/
98	*Phytophthora* Functional Genomics Database	ESTs and expression data from *P.infestans* and *P.sojae*	http://www.pfgd.org/pfgd/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
	5.3.5. Invertebrates		
	5.3.5.1. *Caenorhabditis elegans*		
430	*C.elegans* Project	Genome sequencing data at the Sanger Institute	http://www.sanger.ac.uk/Projects/C_elegans
238	Intronerator	Introns and splicing in *C.elegans* and *C.briggsae*	http://www.cse.ucsc.edu/~kent/intronerator/
570	RNAiDB	RNAi phenotypic analysis of *C.elegans* genes	http://www.rnai.org/
100	WILMA	*C.elegans* annotation database	http://www.came.sbg.ac.at/wilma/
304	WorfDB	*C.elegans* ORFeome	http://worfdb.dfci.harvard.edu/
51	WormBase	Data repository for *C.elegans* and *C.briggsae*: curated genome annotation, genetic and physical maps, pathways	http://www.wormbase.org/
	5.3.5.2. *Drosophila melanogaster*		
71	FlyBase	*Drosophila* sequences and genomic information	http://flybase.bio.indiana.edu/
767	FlyBrain	Database of the *Drosophila* nervous system	http://flybrain.neurobio.arizona.edu
670	FlyMine	Integration of insect genomic and proteomic data	http://www.flymine.org/
467	FlyTrap	*Drosophila* mutants created using GFP protein trap strategy	http://flytrap.med.yale.edu/
471	GadFly	Genome annotation database of *Drosophila*	http://www.fruitfly.org
677	GeniSys	Enhancer- and promoter-inserted mutants of *Drosophila*	http://genisys.kaist.ac.kr:8080/
774	DPDB	*Drosophila* polymorphism database	http://dpdb.uab.es/
449	Drosophila microarray project	Data and tools for *Drosophila* gene expression studies	http://www.flyarrays.com/fruitfly
509	InterActive Fly	*Drosophila* genes and their roles in development	http://sdb.bio.purdue.edu/fly/aimain/1aahome.htm
	5.3.5.3. Other invertebrates		
410	AppaDB	A database on the nematode *Pristionchus pacificus*	http://appadb.eb.tuebingen.mpg.de/
643	BeetleBase	Genome database of the beetle *Tribolium castaneum*	http://www.bioinformatics.ksu.edu/BeetleBase/
649	Ciliate IES-MDS Db	Macro- and micronuclear genes in spirotrichous ciliates	http://oxytricha.princeton.edu/dimorphism/database.htm
434	CnidBase	Cnidarian evolution and gene expression database	http://cnidbase.bu.edu/
543	Nematode.net	Parasitic nematode sequencing project	http://nematode.net/
544	NEMBASE	Nematode sequence and functional data database	http://www.nematodes.org/
726	PPNEMA	Plant-parasitic nematode rRNAs	http://bighost.area.ba.cnr.it/PPNEMA/
743	SilkDB	Silkworm *Bombyx mori* ESTs, mutants, photographs	http://www.ab.a.u-tokyo.ac.jp/genome/
744	SilkSatDb	A microsatellite database of the silkworm *Bombyx mori*	http://www.cdfd.org.in/silksatdb/
747	SpodoBase	Genomics of the butterfly *Spodoptera frugiperda*	http://bioweb.ensam.inra.fr/spodobase/
	6. Metabolic Enzymes and Pathways; Signaling Pathways		
	6.1. Enzymes and Enzyme Nomenclature		
421	BRENDA	Enzyme names and biochemical properties	http://www.brenda.uni-koeln.de
109	ENZYME	Enzyme nomenclature and properties	http://www.expasy.org/enzyme
459	Enzyme Nomenclature	IUBMB Nomenclature Committee recommendations	http://www.chem.qmw.ac.uk/iubmb/enzyme
613	EzCatDB	Enzyme Catalytic Mechanism Database	http://mbs.cbrc.jp/EzCatDB/
508	IntEnz	Integrated enzyme database and enzyme nomenclature	http://www.ebi.ac.uk/intenz
716	PDBrtf	Representation of target families of enzymes in PDB	http://cgl.imim.es/pdbrtf/
758	SCOPEC	Mapping of catalytic function to domain structure	http://www.enzome.com/databases/scopec.php
	6.2. Metabolic Pathways		
644	BioSilico	Integrated access to various metabolic databases	http://biosilico.kaist.ac.kr/
112	KEGG Pathway	Metabolic and regulatory pathways in complete genomes	http://www.genome.jp/kegg/pathway.html
114	MetaCyc	Metabolic pathways and enzymes from various organisms	http://metacyc.org
115	PathDB	Biochemical pathways, compounds and metabolism	http://www.ncgr.org/pathdb
117	UM-BBD	University of Minnesota biocatalysis and biodegradation database	http://umbbd.ahc.umn.edu/
	6.3. Intermolecular Interactions and Signaling Pathways		
633	3DID	3D interacting domains: domain–domain interactions in proteins with known 3D structures	http://3did.embl.de
405	aMAZE	A system for the annotation, management, and analysis of biochemical and signalling pathway networks	http://www.amaze.ulb.ac.be/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
103	BIND	Biomolecular interaction network database	http://www.bind.ca
419	BioCarta	Online maps of metabolic and signaling pathways	http://www.biocarta.com/genes/allPathways.asp
422	BRITE	Biomolecular relations in information transmission and expression, part of KEGG	http://www.genome.ad.jp/brite
659	DDIB	Database of domain interactions and binding	http://www.ddib.org/
104	DIP	Database of interacting proteins: experimentally determined protein–protein interactions	http://dip.doe-mbi.ucla.edu
105	DRC	Database of ribosomal crosslinks	http://www.mpimg-berlin-dahlem.mpg.de/~ag_ribo/ag_brimacombe/drc
329	GeneNet	Database on gene network components	http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet
664	hp-DPI	Database of protein interactions in *Helicobacter pylori*	http://dpi.nhri.org.tw/hp/
688	HPID	Human protein interaction database	http://www.hpid.org/
507	IntAct project	Protein–protein interaction data	http://www.ebi.ac.uk/intact
770	Inter-Chain Beta-Sheets	Protein–protein interactions mediated by interchain beta-sheet formation	http://www.igb.uci.edu/servers/icbs/
510	InterDom	Putative protein domain interactions	http://interdom.lit.org.sg
718	PDZBase	Protein–protein interactions involving PDZ domains	http://icb.med.cornell.edu/services/pdz/start
749	Protein-protein interfaces	Interacting residues in protein–protein interfaces in PDB	http://home.ku.edu.tr/~okeskin/INTERFACE/INTERFACES.html
773	PINdb	Proteins interacting in nucleus (human and yeast)	http://pin.mskcc.org/
748	POINT	Prediction of human protein–protein interactome	http://point.nchc.org.tw/
616	PSIbase	Interaction of proteins with known 3D structures	http://psimap.kaist.ac.kr/
612	Reactome	A knowledgebase of biological pathways	http://www.reactome.org/
571	ROSPath	Reactive oxygen species (ROS) signaling pathway	http://rospath.ewha.ac.kr/
395	STCDB	Signal transductions classification database	http://bibiserv.techfak.uni-bielefeld.de/stcdb/
582	STRING	Predicted functional associations between proteins	http://string.embl.de/
341	TRANSPATH	Gene regulatory networks and microarray analysis	http://www.biobase.de/pages/products/databases.html

7. Human and other Vertebrate Genomes

7.1. Model organisms, comparative genomics

No.[b]	Database name	Full name and/or description	URL
63	ACeDB	*C.elegans*, *S.pombe* and human genomic information	http://www.acedb.org/
26	AllGenes	Human and mouse gene, transcript and protein annotation	http://www.allgenes.org/
65	ArkDB	Genome databases for farm and other animals	http://www.thearkdb.org/
647	ChickVD	Sequence variation in the chicken genome	http://chicken.genomics.org.cn/
286	Cre Transgenic Database	Cre transgenic mouse lines with links to publications	http://www.mshri.on.ca/nagy/
660	DED	Database of evolutionary distances	http://warta.bio.psu.edu/DED/
27	Ensembl	Annotated information on eukaryotic genomes	http://www.ensembl.org/
465	FANTOM	Functional annotation of mouse full-length cDNA clones	http://fantom2.gsc.riken.go.jp
468	FREP	Functional repeats in mouse cDNAs	http://facts.gsc.riken.go.jp/FREP/
673	GALA	Genomic alignment, annotation and experimental results	http://gala.cse.psu.edu/
347	GenetPig	Genes controlling economic traits in pig	http://www.infobiogen.fr/services/Genetpig
605	HomoloGene	Automatically detected homologous genes in complete eukaryotic genomes	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=homologene
690	Inparanoid	A database of eukaryotic orthologs	http://inparanoid.cgb.ki.se/
696	IPI	International protein index: non-redundant sets of human, mouse and rat proteins	http://www.ebi.ac.uk/IPI
777	KaryotypeDB	Karyotype and chromosome information for animal and plant species	http://www.nenno.it/karyotypedb/
400	KOG	Eukaryotic orthologous groups of proteins	http://www.ncbi.nlm.nih.gov/COG/new/shokog.cgi
87	Mouse Genome Informatics	Formerly mouse genome database	http://www.informatics.jax.org/
540	MTID	Mouse transposon insertion database	http://mouse.ccgb.umn.edu/transposon/
703	NegProt	Negative Proteome: a tool for comparison of complete proteomes	http://superfly.ucsd.edu/negprot
556	PEDE	Pig EST data explorer: full-length cDNAs and ESTs	http://pede.gene.staff.or.jp/
665	PhenomicDB	Comparison of phenotypes of orthologous genes in human and model organisms	http://www.phenomicdb.de/
724	Polymorphix	A database of sequence polymorphisms	http://pbil.univ-lyon1.fr/polymorphix/query.php
93	Rat Genome Database	Rat genetic and genomic data	http://rgd.mcw.edu/
625	RatMap	Rat genome tools and data	http://ratmap.org/

Actually header: Nucleic Acids Research, 2005, Vol. 33, Database issue D17

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
751	TAED	The adaptive evolution database: a phylogeny-based tool for comparative genomics	http://www.bioinfo.no/tools/TAED
5	TIGR Gene Indices	Organism-specific databases of EST and gene sequences	http://www.tigr.org/tdb/tgi.shtml
6	UniGene	Unified clusters of ESTs and full-length mRNA sequences	http://www.ncbi.nlm.nih.gov/UniGene/
319	UniSTS	Unified view of sequence tagged sites with mapping data	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unists
783	VEGA	Vertebrate genome annotation: a repository for manual annotation of finished vertebrate genome sequences	http://vega.sanger.ac.uk/
101	ZFIN	Zebrafish information network	http://zfin.org/
	7.2. Human genome databases, maps and viewers		
27	Ensembl	Annotated information on eukaryotic genomes	http://www.ensembl.org/
404	AluGene	Complete *Alu* map in the human genome	http://alugene.tau.ac.il/
349	CroW 21	Human chromosome 21 database	http://bioinfo.weizmann.ac.il/crow21/
55	GB4-RH	Genebridge4 human radiation hybrid maps	http://www.sanger.ac.uk/Software/RHserver/RHserver.shtml
56	GDB	Human genes and genomic maps	http://www.gdb.org/
57	GenAtlas	Human genes, markers and phenotypes	http://www.genatlas.org/
350	GeneCards	Integrated database of human genes, maps, proteins and diseases	http://bioinfo.weizmann.ac.il/cards/
348	GeneLoc	Gene location database (formerly UDB—Unified database for human genome mapping)	http://genecards.weizmann.ac.il/geneloc/
327	GeneNest	Gene indices of human, mouse, zebrafish, etc.	http://genenest.molgen.mpg.de/
59	GenMapDB	Mapped human BAC clones	http://genomics.med.upenn.edu/genmapdb
35	Gene Resource Locator	Alignment of ESTs with finished human sequence	http://grl.gi.k.u-tokyo.ac.jp/
324	HOWDY	Human organized whole genome database	http://www-alis.tokyo.jst.go.jp/HOWDY/
60	HuGeMap	Human genome genetic and physical map data	http://www.infobiogen.fr/services/Hugemap
77	Human BAC Ends	Non-redundant human BAC end sequences	http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_intro.html
597	Human Genome Segmental Duplication Database	Segmental duplications in the human genome	http://projects.tcag.ca/humandup
61	IXDB	Physical maps of human chromosome X	http://ixdb.mpimg-berlin-dahlem.mpg.de/
697	L1Base	Functional annotation and prediction of LINE-1 elements	http://line1.molgen.mpg.de/
54	Map Viewer	Display of genomic information by chromosomal position	http://www.ncbi.nlm.nih.gov/mapview/
600	MGC	Mammalian genome collection: full-length ORFs for human, mouse and rat genes	http://mgc.nci.nih.gov/
391	NCBI RefSeq	Non-redundant collection of naturally occurring biological molecules	http://www.ncbi.nlm.nih.gov/RefSeq/
553	ParaDB	Paralogy mapping in human genomes	http://abi.marseille.inserm.fr/paradb/
62	RHdb	Radiation hybrid map data	http://www.ebi.ac.uk/RHdb
592	SKY/M-FISH and CGH	Fluorescent images of chromosomes and cytogenetic data	http://www.ncbi.nlm.nih.gov/sky/
4	STACK	Sequence tag alignment and consensus knowledgebase	http://www.sanbi.ac.za/Dbases.html
596	The Chromosome 7 Annotation Project	A comprehensive description of human chromosome 7	http://www.chr7.org/
684	TRBase	Tandem repeats in the human genome	http://bioinfo.ex.ac.uk/trbase
316	UCSC Genome Browser	Genome assemblies and annotation	http://genome.ucsc.edu/
	7.3. Human proteins		
685	H-InvDB	Full-length human cDNA clones	http://www.h-invitational.jp/
499	HPMR	Human plasma membrane receptome: sequences, literature and expression data	http://receptome.stanford.edu/
500	HPRD	Human protein reference database: domain architecture, post-translational modifications and disease association	http://www.hprd.org
37	HUNT	Human novel transcripts: annotated full-length cDNAs	http://www.hri.co.jp/HUNT
171	HUGE	Human unidentified gene-encoded large (>50 kDa) protein and cDNA sequences	http://www.kazusa.or.jp/huge
522	LIFEdb	Localization, interaction and functions of human proteins	http://www.dkfz.de/LIFEdb
312	trome, trEST and trGEN:	Databases of predicted human protein sequences	ftp://ftp.isrec.isb-sib.ch/pub/databases/
	8. Human Genes and Diseases		
	8.1. General Databases		
661	DG-CST	Disease gene conserved sequence tags	http://143.225.208.11/cst3/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
683	HCAD	Human chromosome aberration database: chromosomal breakpoints and affected genes	http://www.pdg.cnb.uam.es/UniPub/HCAD/
8	Homophila	*Drosophila* homologs of human disease genes to	http://superfly.ucsd.edu/homophila/
548	OMIA	Online Mendelian inheritance in animals: a catalog of animal genetic and genomic disorders	http://www.angis.org.au/omia
143	OMIM	Online Mendelian inheritance in man: a catalog of human genetic and genomic disorders	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
550	ORFDB	Collection of ORFs that are sold by Invitrogen	http://orf.invitrogen.com/
554	PathBase	European mutant mice histopathology database: images	http://www.pathbase.net/
146	PMD	Compilation of protein mutant data	http://pmd.ddbj.nig.ac.jp/
358	SOURCE	Functional genomics resource for human, mouse and rat	http://source.stanford.edu/

8.2. Human Mutations Databases
8.2.1. General polymorphism databases

No.[b]	Database name	Full name and/or description	URL
119	ALFRED	Allele frequencies and DNA polymorphisms	http://alfred.med.yale.edu/
416	BayGenomics	Genes relevant to cardiovascular and pulmonary disease	http://baygenomics.ucsf.edu/
654	Cypriot national mutation database	Disease mutations in the Cypriot population	http://www.goldenhelix.org/cypriot/
655	Database of Genomic Variants	Human genomic variants: frequency, segmental duplications and genome assembly gaps	http://projects.tcag.ca/variation/
595	dbQSNP	Quantification of SNP allele frequencies database	http://qsnp.gen.kyushu-u.ac.jp/
127	dbSNP	Database of single nucleotide polymorphisms	http://www.ncbi.nlm.nih.gov/SNP/
669	FESD	Functional element SNPs database: SNPs located within promoters, UTRs, etc., of human genes	http://combio.kribb.re.kr/ksnp/resd/
496	HGVS Databases	A compilation of human mutation databases	http://www.hgvs.org/
131	HGVbase	Human genome variation database: curated human polymorphisms	http://hgvbase.cgb.ki.se/
133	HGMD	Human gene mutation database	http://www.hgmd.org/
367	IPD	Immuno polymorphism database	http://www.ebi.ac.uk/ipd
517	JSNP	Japanese SNP database	http://snp.ims.u-tokyo.ac.jp/
45	rSNP Guide	SNPs in regulatory gene regions	http://wwwmgs.bionet.nsc.ru/mgs/systems/rsnp/
344	SNP Consortium database	SNP Consortium data	http://snp.cshl.org/
626	SNPeffect	Phenotypic effects of human coding SNPs	http://snpeffect.vib.be/
590	TopoSNP	Topographic database of non-synonymous SNPs	http://gila.bioengr.uic.edu/snp/toposnp
755	TPMD	Taiwan polymorphic microsatellite marker database	http://tpmd.nhri.org.tw/

8.2.2. Cancer

No.[b]	Database name	Full name and/or description	URL
122	Atlas of Genetics and Cytogenetics in Oncology and Haematology	Cancer-related genes, chromosomal abnormalities in oncology and haematology, and cancer-prone diseases	http://www.infobiogen.fr/services/chromcancer/
593	Cancer Chromosomes	Cytogenetic, clinical and reference information on cancer-related aberrations	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=cancerchromosomes
431	CGED	Cancer gene expression database	http://love2.aist-nara.ac.jp/CGED
651	COSMIC	Catalogue of somatic mutations in cancer: sequence data, samples and publications	http://www.sanger.ac.uk/perl/CGP/cosmic
126	Germline p53 Mutations	Mutations in human tumor and cell line p53 gene	http://www.lf2.cuni.cz/win/projects/germline_mut_p53.htm
362	IARC TP53 Database	Human TP53 somatic and germline mutations	http://www.iarc.fr/p53/
152	MTB	Mouse tumor biology database: tumor types, genes, classification, incidence, pathology	http://tumor.informatics.jax.org/
709	OncoMine	Cancer microarray data by gene or cancer type	http://www.oncomine.org/
153	Oral Cancer Gene Database	Cellular and molecular data for genes involved in oral cancer	http://www.tumor-gene.org/Oral/oral.html
148	RB1 Gene Mutation DB	Mutations in the human retinoblastoma (RB1) gene	http://www.d-lohmann.de/Rb/
574	RTCGD	Mouse retroviral tagged cancer gene database	http://rtcgd.ncifcrf.gov/
579	SNP500Cancer	Re-sequenced SNPs from 102 reference samples	http://snp500cancer.nci.nih.gov
149	SV40 Large T-Antigen Mutants	Mutations in SV40 large tumor antigen gene	http://supernova.bio.pitt.edu/pipaslab/
155	Tumor Gene Family Databases	Cellular, molecular and biological data about genes involved in various cancers	http://www.tumor-gene.org/tgdf.html

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
	8.2.3. Gene-, system- or disease-specific		
768	ALPSbase	Autoimmune lymphoproliferative syndrome database	http://research.nhgri.nih.gov/alps/
120	Androgen Receptor Gene Mutations Database	Mutations in the androgen receptor gene	http://www.mcgill.ca/androgendb/
123	BTKbase	Mutation registry for X-linked agammaglobulinemia	http://bioinf.uta.fi/BTKbase/
594	CarpeDB	Comprehensive database on the genetics of epilepsy	http://www.carpedb.ua.edu/
124	CASRDB	Calcium-sensing receptor database: CASR mutations causing hypercalcemia and/or hyperparathyroidism	http://www.casrdb.mcgill.ca/
125	Cytokine Gene Polymorphism in Human Disease	Cytokine gene polymorphism literature database	http://bris.ac.uk/pathandmicro/services/GAI/cytokine4.htm
137	Collagen Mutation Database	Human type I and type III collagen gene mutations	http://www.le.ac.uk/genetics/collagen/
460	ERGDB	Estrogen responsive genes database	http://research.i2r.a-star.edu.sg/promoter/Ergdb-v11/
164	FUNPEP	Low-complexity peptides capable of forming amyloid plaque	http://www.cmbi.kun.nl/swift/FUNPEP/gergo/
363	GOLD.db	Genomics of lipid-associated disorders database	http://gold.tugraz.at
129	GRAP	Mutants of G-protein coupled receptors of family A	http://tinygrap.uit.no/GRAP/
130	HaemB	Factor IX gene mutations, insertions and deletions	http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html
491	HbVar	Human hemoglobin variants and thalassemias	http://globin.cse.psu.edu/globin/hbvar
680	HAGR	Human ageing genomic resources: genes related to ageing in humans and model organisms	http://genomics.senescence.info/
134	Human p53/hprt, rodent lacI/lacZ databases	Mutations at the human p53 and hprt genes; rodent transgenic lacI and lacZ mutations	http://www.ibiblio.org/dnam/mainpage.html
135	Human PAX2 Allelic Variant Database	Mutations in human PAX2 gene	http://pax2.hgu.mrc.ac.uk/
136	Human PAX6 Allelic Variant Database	Mutations in human PAX6 gene	http://pax6.hgu.mrc.ac.uk/
506	INFEVERS	Hereditary inflammatory disorder and familial mediterranean fever mutation data	http://fmf.igh.cnrs.fr/infevers
139	KinMutBase	Disease-causing protein kinase mutations	http://www.uta.fi/imt/bioinfo/KinMutBase/
523	Lowe Syndrome Mutation Database	Mutations causing Lowe oculocerebrorenal syndrome	http://research.nhgri.nih.gov/lowe/
142	NCL Mutation Database	Polymorphisms in neuronal ceroid lipofuscinoses genes	http://www.ucl.ac.uk/ncl/
144	PAHdb	Mutations at the phenylalanine hydroxylase locus	http://www.mcgill.ca/pahdb/
559	PGDB	Prostate and prostatic diseases gene database	http://www.ucsf.edu/pgdb/
145	PHEXdb	PHEX mutations causing X-linked hypophosphatemia	http://www.phexdb.mcgill.ca/
147	PTCH1 Mutation Database	Mutations and SNPs found in PTCH1 gene	http://www.cybergene.se/PTCH/ptchbase.html
632	SCAdb	Spinocerebellar ataxia candidate gene database	http://ymbc.ym.edu.tw/cgi-bin/SCA/list.cgi?display=map
752	T1Dbase	A resource for type 1 diabetes research	http://t1dbase.org/
	The Autism Chromosome Rearrangement Database	Curated collection of genomic features related to autism research	http://projects.tcag.ca/autism
753	The Lafora Database	Mutations and polymorphisms associated with Lafora progressive myoclonus epilepsy	http://projects.tcag.ca/lafora/
	9. Microarray Data and other Gene Expression Databases		
634	5'SAGE	5'-end serial analysis of gene expression	http://5sage.gi.k.u-tokyo.ac.jp/
338	ArrayExpress	Public collection of microarray gene expression data	http://www.ebi.ac.uk/arrayexpress
11	Axeldb	Gene expression in *Xenopus laevis*	http://www.dkfz-heidelberg.de/abt0135/axeldb.htm
12	BodyMap	Human and mouse gene expression data	http://bodymap.ims.u-tokyo.ac.jp/
417	BGED	Brain gene expression database	http://love2.aist-nara.ac.jp/BGED
432	CleanEx	Expression reference database, linking heterogeneous expression data to facilitate cross-dataset comparisons	http://www.cleanex.isb-sib.ch/
657	dbERGEII	Database of experimental results on gene expression: genomic alignment, annotation and experimental data	http://dberge.cse.psu.edu/menu.html
454	EICO DB	Expression-based imprint candidate organiser: a database for discovery of novel imprinted genes	http://fantom2.gsc.riken.jp/EICODB/
455	*emap* Atlas	Edinburgh mouse atlas: a digital atlas of mouse embryo development and spatially mapped gene expression	http://genex.hgu.mrc.ac.uk/

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
13	EPConDB	Endocrine pancreas consortium database	http://www.cbil.upenn.edu/EPConDB
110	EpoDB	Genes expressed during human erythropoiesis	http://www.cbil.upenn.edu/EpoDB/
14	FlyView	*Drosophila* development and genetics	http://pbio07.uni-muenster.de/
326	GeneAnnot	Revised annotation of Affymetrix human gene probe sets	http://genecards.weizmann.ac.il/geneannot/
325	GeneNote	Human genes expression profiles in healthy tissues	http://genecards.weizmann.ac.il/genenote/
330	GenePaint	Gene expression patterns in the mouse	http://www.genepaint.org/Frameset.html
676	GeneTide	A transcriptome-focused member of the GeneCards suite	http://genecards.weizmann.ac.il/genetide/
481	GeneTrap	Expression patterns in an embryonic stem library of gene trap insertions	http://www.cmhd.ca/sub/genetrap.asp
603	GEO	Gene expression omnibus: gene expression profiles	http://www.ncbi.nlm.nih.gov/geo/
485	GermOnline	Gene expression in mitotic and meiotic cell cycle	http://www.germonline.org/
15	GXD	Mouse gene expression database	http://www.informatics.jax.org/menus/expression_menu.shtml
681	H-ANGEL	Human anatomic gene expression library	http://www.jbirc.aist.go.jp/hinv/index.jsp
493	HemBase	Genes expressed in differentiating human erythroid cells	http://hembase.niddk.nih.gov/
23	HugeIndex	Expression levels of human genes in normal tissues	http://hugeindex.org/
17	Kidney Development Database	Kidney development and gene expression	http://golgi.ana.ed.ac.uk/kidhome.html
778	LOLA	List of lists annotated: a comparison of gene sets identified in different microarray experiments	http://www.lola.gwu.edu/
18	MAGEST	Ascidian (*Halocynthia roretzi*) gene expression patterns	http://www.genome.ad.jp/magest
699	MAMEP	Molecular anatomy of the mouse embryo project: gene expression data on mouse embryos	http://mamep.molgen.mpg.de/
339	MEPD	Medaka (freshwater fish *Oryzias latipes*) gene expression pattern database	http://www.embl.de/mepd/
19	MethDB	DNA methylation data, patterns and profiles	http://www.methdb.de/
537	Mouse SAGE	SAGE libraries from various mouse tissues and cell lines	http://mouse.biomed.cas.cz/sage
541	NASCarrays	Nottingham *Arabidopsis* Stock Centre microarray database	http://affymetrix.arabidopsis.info
545	NetAffx	Public Affymetrix probesets and annotations	http://www.affymetrix.com/
711	Osteo-Promoter Database	Genes in osteogenic proliferation and differentiation	http://www.opd.tau.ac.il
154	PEDB	Prostate expression database: ESTs from prostate tissue and cell type-specific cDNA libraries	http://www.pedb.org/
558	PEPR	Public expression profiling resource: expression profiles in a variety of diseases and conditions	http://microarray.cnmcresearch.org/pgadatatable.asp
21	RECODE	Genes using programmed translational recoding in their expression	http://recode.genetics.utah.edu/
568	RefExA	Reference database for human gene expression analysis	http://www.lsbm.org/db/index_e.html
739	rOGED	Rat ovarian gene expression database	http://web5.mccs.uky.edu/kolab/rogedendo.aspx
712	SAGEmap	NCBI's resource for SAGE data from various organisms	http://www.ncbi.nlm.nih.gov/SAGE
742	SIEGE	Smoking Induced Epithelial Gene Expression	http://pulm.bumc.bu.edu/siegeDB
22	Stanford Microarray Database	Raw and normalized data from microarray experiments	http://genome-www.stanford.edu/microarray
24	Tooth Development Database	Gene expression in dental tissue	http://bite-it.helsinki.fi/
10. Proteomics Resources			
606	2D-PAGE	Proteome database system for microbial research	http://www.mpiib-berlin.mpg.de/2D-PAGE
731	DynaProt 2D	Proteome database of *Lactococcus lactis*	http://www.wzw.tum.de/proteomik/lactis/
222	GelBank	2D gel electrophoresis patterns of proteins from complete microbial genomes	http://gelbank.anl.gov/
710	Open Proteomics Database	Mass-spectrometry-based proteomics data for human, yeast, *E.coli* and *Mycobacterium*	http://bioinformatics.icmb.utexas.edu/OPD/
377	PEP	Predictions for entire proteomes: summarized analyses of protein sequences	http://cubic.bioc.columbia.edu/pep/
281	RESID	Pre-, co- and post-translational protein modifications	http://www-nbrf.georgetown.edu/pirwww/dbinfo/resid.html
225	SWISS-2DPAGE	Annotated 2D gel electrophoresis database	http://www.expasy.org/ch2d/
11. Other Molecular Biology Databases			
	11.1. Drugs and drug design		
407	ANTIMIC	Database of natural antimicrobial peptides	http://research.i2r.a-star.edu.sg/Templar/DB/ANTIMIC/
781	AOBase	Antisense oligonucleotide selection and design	http://www.bio-inf.net/aobase/
408	APD	Antimicrobial peptide database	http://aps.unmc.edu/AP/main.php

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
423	BSD	Biodegradative strain database: microorganisms that can degrade aromatic and other organic compounds	http://bsd.cme.msu.edu/
443	DART	Drug adverse reaction target database	http://xin.cz3.nus.edu.sg/group/drt/dart.asp
701	MetaRouter	Compounds and pathways related to bioremediation	http://pdg.cnb.uam.es/MetaRouter/
186	Peptaibol	Peptaibol (antibiotic peptide) sequences	http://www.cryst.bbk.ac.uk/peptaibol/welcome.html
392	PharmGKB	Pharmacogenomics knowledge base: effect of genetic variation on drug responses	http://www.pharmgkb.org/
315	TTD	Therapeutic target database	http://xin.cz3.nus.edu.sg/group/cjttd/ttd.asp
11.2. Probes			
505	IMGT/PRIMER-DB	Immunogenetics oligonucleotide primer database	http://imgt3d.igh.cnrs.fr/PrimerDB/Query_PrDB.pl
296	MPDB	Synthetic oligonucleotides useful as primers or probes	http://www.biotech.ist.unige.it/interlab/mpdb.html
728	PrimerPCR	PCR primers for eukaryotic and prokaryotic genes	http://bioinfo.ebc.ee/PrimerStudio/
390	probeBase	rRNA-targeted oligonucleotide probe sequences, DNA microarray layouts and associated information	http://www.microbial-ecology.net/probebase
736	QPPD	Quantitative PCR Primer Database for human and mouse	http://web.ncifcrf.gov/rtp/GEL/primerdb/default.asp
356	RTPrimerDB	Real-time PCR primer and probe sequences	http://medgen.ugent.be/rtprimerdb/
11.3. Unclassified databases			
298	PubMed	Citations and abstracts of biomedical literature	http://pubmed.gov/
256	BioImage	Database of multidimensional biological images	http://www.bioimage.org/
12. Organelle Databases			
74	GOBASE	Organelle genome database	http://megasun.bch.umontreal.ca/gobase/
547	OGRe	Organelle genome retrieval system	http://ogre.mcmaster.ca/
601	Organelle genomes	NCBI's organelle genome resource	http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/organelles.html
722	PLprot	*Arabidopsis thaliana* chloroplast protein database	http://www.pb.ipw.biol.ethz.ch/proteomics
713	Organelle DB	Organelle proteins and subcellular structures	http://organelledb.lsi.umich.edu/
12.1. Mitochondrial Genes and Proteins			
637	AMPDB	*Arabidopsis* mitochondrial protein database	http://www.mitoz.bcs.uwa.edu.au/AMPDB/
686	HMPD	Human mitochondrial protein database	http://bioinfo.nist.gov:8080/examples/servlets/index.html
38	HvrBase	Primate mitochondrial DNA control region sequences	http://www.hvrbase.org/
64	Mitochondriome	Metazoan mitochondrial genes	http://www.ba.itb.cnr.it/mitochondriome
83	MitoDat	Mitochondrial proteins (predominantly human)	http://www-lecb.ncifcrf.gov/mitoDat/
226	MitoDrome	Nuclear-encoded mitochondrial proteins of *Drosophila*	http://www2.ba.itb.cnr.it/MitoDrome/
84	MitoMap	Human mitochondrial genome	http://www.mitomap.org/
85	MitoNuc	Nuclear genes coding for mitochondrial proteins	http://www2.ba.itb.cnr.it/MitoNuc/
86	MITOP2	Mitochondrial proteins, genes and diseases	http://ihg.gsf.de/mitop2/start.jsp
531	MitoPD	Yeast mitochondrial protein database	http://bmerc-www.bu.edu/mito
532	*MitoProteome*	Experimentally described human mitochondrial proteins	http://www.mitoproteome.org
538	MPIMP	Mitochondrial protein import machinery of plants	http://millar3.biochem.uwa.edu.au/~lister/index.html
241	PLMItRNA	Plant mitochondrial tRNA	http://bighost.area.ba.cnr.it/PLMItRNA/
13. Plant Databases			
13.1. General plant databases			
599	BarleyBase	Expression profiling of plant genomes	http://www.barleybase.org/
624	CR-EST	Crop ESTs: barley, pea, wheat and potato	http://pgrc.ipk-gatersleben.de/cr-est/
67	CropNet	Genome mapping in crop plants	http://ukcrop.net/
128	FLAGdb++	Integrative database about plant genomes	http://genoplante-info.infobiogen.fr/FLAGdb/
351	GénoPlante-Info	Plant genomic data from the Génoplante consortium	http://genoplante-info.infobiogen.fr/
488	GrainGenes	Genes and phenotypes of wheat, barley, rye, triticale, oats	http://wheat.pw.usda.gov or http://www.graingenes.org
607	Gramene	A resource for comparative grass genomics	http://www.gramene.org/
81	Mendel	Annotated plant ESTs and STSs	http://www.mendel.ac.uk/
581	openSputnik	Plant EST clustering and functional annotation	http://sputnik.btk.fi

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
560	PhytoProt	Clusters of (predicted) plant proteins	http://genoplante-info.infobiogen.fr/phytoprot
721	PlantMarkers	A database of predicted molecular markers from plants	http://markers.btk.fi/
355	PlantGDB	Plant genome database: actively transcribed plant genes	http://www.plantgdb.org/
189	PLANT-PIs	Plant protease inhibitors	http://bighost.area.ba.cnr.it/PLANT-PIs
371	PlantsP/PlantsT	Plant proteins involved in phosphorylation and transport	http://plantsp.sdsc.edu/
588	TIGR plant repeat database	Classification of repetitive sequences in plant genomes	http://www.tigr.org/tdb/e2k1/plant.repeats
313	TropGENE DB	Genes and genomes of sugarcane, banana, cocoa	http://tropgenedb.cirad.fr/
	13.2. *Arabidopsis thaliana*		
636	AGNS	*Arabidopsis* GeneNet supplementary: gene expression and phenotypes of mutants and transgens	http://emj-pc.ics.uci.edu/mgs/dbases/agns
618	AGRIS	*Arabidopsis* gene regulatory information server: promoters, transcription factors and their target genes	http://arabidopsis.med.ohio-state.edu/
780	*Arabidopsis* MPSS	*Arabidopsis* gene expression detected by massively parallel signature sequencing	http://mpss.udel.edu/at/
638	*Arabidopsis* Nucleolar Protein Database	Comparative analysis of human and *Arabidopsis* nucleolar proteomes	http://bioinf.scri.sari.ac.uk/cgi-bin/atnopdb/proteome_comparison
640	ASRP	*Arabidopsis thaliana* small RNA project	http://cgrb.orst.edu/smallRNA/db/
412	ARAMEMNON	*Arabidopsis thaliana* membrane proteins and transporters	http://aramemnon.botanik.uni-koeln.de/
765	AthaMap	Genome-wide map of putative transcription factor binding sites in *Arabidopsis thaliana*	http://www.athamap.de/
427	CATMA	Complete *Arabidopsis* transcriptome microarray	http://www.catma.org
656	DATF	Database of *Arabidopsis* transcription factors	http://datf.cbi.pku.edu.cn/
672	GabiPD		http://gabi.rzpd.de/
675	GeneFarm	Expert annotation of *Arabidopsis* gene and protein families	http://genoplante-info.infobiogen.fr/Genefarm/
527	MAtDB	MIPS Arabidopsis thaliana database	http://mips.gsf.de/proj/thal/db
738	RARGE	RIKEN *Arabidopsis* genome encyclopedia: cDNAs, mutants and microarray data	http://rarge.gsc.riken.jp/
578	SeedGenes	Genes essential for *Arabidopsis* development	http://www.seedgenes.org/
97	TAIR	The *Arabidopsis* information resource	http://www.arabidopsis.org/
627	WAtDB	Wageningen *Arabidopsis thaliana* database: mutants, transgenic lines and natural variants	http://www.watdb.nl/
	13.3. Rice		
418	BGI-RISe	Beijing genomics institute rice information system	http://rise.genomics.org.cn/
79	INE	Integrated rice genome explorer	http://rgp.dna.affrc.go.jp/giot/INE.html
353	IRIS	International rice information system	http://www.iris.irri.org/
536	MOsDB	MIPS *Oryza sativa* database	http://mips.gsf.de/proj/plant/jsf/rice/index.jsp
90	Oryzabase	Rice genetics and genomics	http://www.shigen.nig.ac.jp/rice/oryzabase/
628	Oryza Tag Line database	T-DNA insertion mutants of rice	http://genoplante-info.infobiogen.fr/OryzaTagLine/
737	RAD	Rice annotation database	http://golgi.gs.dna.affrc.go.jp/SY-1102/rad/index.html
336	RiceGAAS	Rice genome automated annotation system	http://ricegaas.dna.affrc.go.jp/
569	Rice PIPELINE	Unification tool for rice databases	http://cdna01.dna.affrc.go.jp/PIPE
572	Rice proteome database	Rice proteome database	http://gene64.dna.affrc.go.jp/RPD/main_en.html
	13.4. Other plants		
610	*Brassica* ASTRA	A database for *Brassica* genomic research	http://hornbill.cspp.latrobe.edu.au/cgi-binpub/index.pl
526	MaizeGDB	Maize genetics and genomics database	http://www.maizegdb.org/
80	LIS (formerly MGI)	Legume information server (formerly Medicago genome initiative): ESTs, gene expression and proteomic data	http://www.comparative-legumes.org/
539	MtDB	*Medicago trunculata* genome database	http://www.medicago.org/MtDB
620	PoMaMo	Potato Maps and More: Potato genome data	https://gabi.rzpd.de/PoMaMo.html
766	SGMD	Soybean genomics and microarray database	http://psi081.ba.ars.usda.gov/SGMD/default.htm

Table 1. *Continued*

No.[b]	Database name	Full name and/or description	URL
	14. Immunological Databases		
642	BCIpep	A database of B-cell epitopes	http://bioinformatics.uams.edu/mirror/bcipep/
604	dbMHC	Genetic and clinical database of the human MHC	http://www.ncbi.nlm.nih.gov/mhc/
150	FIMM	Functional molecular immunology data	http://research.i2r.a-star.edu.sg/fimm
682	Haptendb	Curated database of hapten molecules	http://www.imtech.res.in/raghava/haptendb/
779	HLA Ligand/Motif	A database and search tool for HLA sequences	http://hlaligand.ouhsc.edu/
501	IL2Rgbase	X-linked severe combined immunodeficiency mutations	http://research.nhgri.nih.gov/scid/
172	IMGT	International immunogenetics information system: immunoglobulins, T cell receptors, MHC and RPI	http://imgt.cines.fr/
503	IMGT/Gene-DB	Vertebrate immunoglobulin and T cell receptor genes	http://imgt.cines.fr/cgi-bin/GENElect.jv
173	IMGT/HLA	Polymorphism of human MHC and related genes	http://www.ebi.ac.uk/imgt/hla/
504	IMGT/LIGM-DB	Immunoglobulin, T cell receptor and MHC nucleotide sequences from human and other vertebrates	http://imgt.cines.fr/cgi-bin/IMGTlect.jv
16	Interferon Stimulated Gene Database	Genes induced by treatment with interferons	http://www.lerner.ccf.org/labs/williams/xchip-html.cgi
692	IPD-ESTDAB	Immunologically characterized melanoma cell lines	http://www.ebi.ac.uk/ipd/estdab/
693	IPD-HPA	Immuno polymorphism of human platelet antigens	http://www.ebi.ac.uk/ipd/hpa/
694	IPD-KIR	Immuno polymorphism of killer-cell Ig-like receptors	http://www.ebi.ac.uk/ipd/kir/
691	IPD-MHC	Sequences of the major histocompatibility complex	http://www.ebi.ac.uk/ipd/mhc/
361	JenPep	Quantitative binding data for immunological protein–peptide interactions	http://www.jenner.ac.uk/Jenpep
702	MHCBN	A database of MHC binding and non-binding peptides	http://www.imtech.res.in/raghava/mhcbn/
181	MHCPEP	MHC-binding peptides	http://wehih.wehi.edu.au/mhcpep/
107	MPID	MHC—peptide interaction database	http://surya.bic.nus.edu.sg/mpid
621	VBASE2	Variable genes from the Ig loci of human and mouse	http://www.vbase2.org

[a]Each database is shown in the list only once, often in a category that was arbitrarily chosen among two or three appropriate ones. In the online version of this list at the NAR website (http://nar.oupjournals.org/), a database can be listed under two categories.
[b]Accession number of the database in the online list; can be used to view the database summary, e.g. http://www3.oup.co.uk/nar/database/summary/1 shows the summary for DDBJ.

to survive, even if they have to change their funding scheme or migrate from one host institution to another. This means that the open database movement is here to stay, and more and more people in the community (as well as in the financing bodies) now appreciate the importance of open databases in spreading knowledge. It is worth noting that the majority of database authors and curators receive little or no remuneration for their efforts and that it is still difficult to obtain money for creating and maintaining a biological database. However, disk space is relatively cheap these days and database maintenance tools are fairly straightforward, so that a decent database can be created on a shoestring budget, often by a graduate student or as a result of a postdoctoral project. Many databases in this compilation originated just that way—as collections of data on a certain research topic that a particular lab was studying anyway, formatted in a user-friendly way by graduate or even undergraduate students as part of their dissertations or course work. Subsequent maintenance and further development of these databases, however, require a commitment that can only be applauded. For scientists from China, France, Japan, Russia and many other countries, making their databases available to the worldwide community also means maintaining them in English, the *lingua franca* of science, which does not always come easily. Such efforts deserve special appreciation.

Speaking of appreciation, those who maintain databases often do not get much credit for their work either. Other than publication in the Database Issue of *Nucleic Acids Research* or in *Bioinformatics*, or an occasional publication in some other journal, there is currently no straightforward way to announce progress. The online summaries published by the database maintainers on the NAR web site partially fill this void. Since the entry number assigned to each database (Table 1) will be stable, these updates can be cited in just the same way as any other online resource. At this time, I would suggest the following format for citing these summaries: 'The ooTFD database (11) is listed with Accession No. 185 in the NAR Molecular Biology Database compilation (1); see the recent summary at http://www3.oup.co.uk/nar/database/summary/185'. Suggestions for a better format are certainly welcome. Suggestions for the inclusion of additional databases in this Collection, as well as for improvements to the category structure, are also encouraged and should be directed to the author at galperin@ncbi.nlm.nih.gov.

REFERENCES

1. Galperin,M.Y. (2004) The Molecular Biology Database Collection: 2004 update. *Nucleic Acids Res.*, **32**, D3–D22.
2. Barta,E., Sebestyén,E., Pálfy,T.B., Tóth,G., Ortutay,C.P. and Patthy,L. (2005) DoOP: Databases of Orthologous Promoters, collections of clusters of orthologous upstream sequences from chordates and plants. *Nucleic Acids Res.*, **33**, D86–D90.
3. Tusnády,G.E., Dosztányi,Z. and Simon,I. (2005) PDB_TM: selection and membrane localization of transmembrane proteins in the Protein Data Bank. *Nucleic Acids Res.*, **33**, D275–D278.
4. Huang,H.-D., Horng,J.-T., Lin,F.-M., Chang,Y.-C. and Huang,C.-C. (2005) SpliceInfo: an information repository for the modes of mRNA alternative splicing in human genome. *Nucleic Acids Res.*, **33**, D80–D85.
5. Chang,Y.-H., Su,W.-H., Lee,T.-C., Sun,H.-F.S., Chen,C.-H., Pan,W.-H., Tsai,S.-F. and Jou,Y.-S. (2005) TPMD: a database and resources of microsatellite marker genotyped in Taiwanese populations. *Nucleic Acids Res.*, **33**, D174–D177.
6. Siew,N., Azaria,Y. and Fischer,D. (2004) The ORFanage: an ORFan database. *Nucleic Acids Res.*, **32**, D281–D283.
7. Tanino,M., Debily,M.-A., Tamura,T., Hishiki,T., Ogasawara,O., Murakawa,K., Kawamoto,S., Itoh,K., Watanabe,S., José de Souza,S., Imbeaud,S., Graudens,E., Eveno,E., Hilton,P., Sudo,Y., Kelso,J., Ikeo,K., Imanishi,T., Gojobori,T., Auffray,C., Hide,W. and Okubo,K. (2005) The Human Anatomic Gene Expression Library (H-ANGEL), the H-Inv integrative display of human gene expression across disparate technologies and platforms. *Nucleic Acids Res.*, **33**, D567–D572.
8. Fernandez-Ricaud,L., Warringer,J., Ericson,E., Pylvänäinen,I., Kemp,G., Nerman,O. and Blomberg,A. (2005) PROPHECY – a database for high-resolution phenomics. *Nucleic Acids Res.*, **33**, D369–D373.
9. Whelan,S., de Bakker,P.I. and Goldman,N. (2003) Pandit: a database of protein and associated nucleotide domains with inferred trees. *Bioinformatics*, **19**, 1556–1563.
10. Shah,V., Sridhar,S., Beane,J., Brody,J.S. and Spira,A. (2005) SIEGE: Smoking Induced Epithelial Gene Expression Database. *Nucleic Acids Res.*, **33**, D573–D579.
11. Ghosh,D. (2000) Object-oriented transcription factors database (ooTFD). *Nucleic Acids Res.*, **28**, 308–310.